DOCUMENT RESUME

ED 206 655 TM 810 547

AUTHOR	Taylor, Cie; White, Karl R.
TITLE	Effects of Reinforcement and Training on Title I Students' Group Standardized Test Performance.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C.
PUB DATE	Apr 81
GRANT	BEH-13-443A
NOTE	10p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Academic Achievement; Analysis of Variance; Compensatory Education; Grade 2; *Motivation; Primary Education; *Reading Tests; *Reinforcement; *Standardized Tests; *Testing; Test Validity; *Test Wiseness
IDENTIFIERS	Elementary Secondary Education Act Title I; Stanford Achievement Tests

ABSTRACT

To examine the influence of test wiseness and motivation on Title I students' performance in group standardized tests, the effects of training students and teachers in testing skills and reinforcing student testing behavior were investigated during the administration of the Stanford Achievement Test, Level II, to 597 students. Students were randomly assigned to a combination of the following treatment conditions: (1) provision of external reinforcement, contingent on an increase in the student's reading test percentile rank; (2) teacher training in test administration procedures; and (3) student training in test taking procedures. A three way analysis of variance was computed for the three dependent variables of reading test scores, and student and teacher behavior during testing. Major findings showed that the reading test scores of students in either the reinforced or trained test administrator conditions were significantly higher than the scores of students under nontreatment conditions. There were no significant differences in reading test scores between students who were trained in test wiseness and those who were not trained. Trained teachers exhibited a significantly higher percent of on-task behavior during testing than untrained teachers. There were no significant differences in student on-task behavior under any treatment. (Author/AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED206655

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Effects of Reinforcement and Training on Title I Students' Group Standardized Test Performance

Cie Taylor
Powell Associates, Inc.
Austin, Texas

Karl R. White
Utah State University
Logan, Utah

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

C. Taylor

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM 810 547

A paper presented at the annual meeting of the
American Educational Research Association,
Los Angeles, California
April 13-17, 1981

Executive Summary

Group standardized tests are frequently used to measure achievement of Title I students, although these students often do not understand the testing purpose, the directions, or the question/response formats and may be unmotivated to try their best. To examine the influence of test wiseness and motivation on group standardized test performance, the effects of training students and teachers in testing skills and reinforcing student testing behavior were investigated.

Three treatment conditions were examined during the spring administration of the Stanford Achievement Test, Level II: (a) providing external reinforcement (money) contingent on an increase of student reading test percentile rank, (b) training students in appropriate test taking procedures, and (c) training teachers in appropriate test administration procedures. Twenty-four second grade classrooms (597 students) in eight Title I or similar schools were randomly assigned to each combination of the eight experimental factors generated by the treatment and non-treatment conditions (2 X 2 X 2).

Three dependent variables were examined for differences between groups: (a) reading test score, (b) student on-task behavior during testing, and (c) teacher on-task behavior during testing. A three-way analysis of variance was computed for each dependent variable.

Major findings showed that the reading test scores of students in either the reinforced or trained test administrator conditions were significantly higher than the scores of students under nontreatment conditions. There were no significant differences in reading test scores between students who were trained in test wiseness and those who were not trained. Trained teachers exhibited a significantly higher percent of on-task behavior during testing than untrained teachers. There were no significant differences in student on-task behavior under any treatment.

It was concluded that group achievement test scores from second grade Title I students who are highly motivated or who take tests from teachers trained in proper procedures tend to be more accurate in estimating student knowledge than scores from students not under those conditions.

Effects of Reinforcement and Training on Title I Students' Group Standardized Test Performance[1]

Testing is a major industry in American education. For Title I, testing takes on added importance because standardized tests are the primary tools used to select target students, to make placement and curriculum decisions, and to evaluate programs.

There is some evidence that data typically collected from group administered achievement tests may not be valid indicators of student achievement. Data from informal observations demonstrate that Title I students frequently do not understand the purpose of standardized tests, are unmotivated to do their best, and have difficulty staying on-task and understanding and assimilating the instructions which are given by the test administrator immediately prior to test taking. These factors may lead to anxiety which promotes negative attitudes toward testing and an increased number of errors (Piersel, Brody, & Kratochwill, 1977). Test results from a student who does not understand the reason or procedures for test taking or who is unmotivated to obtain the best possible score have been found to be an underestimation of that student's true level of achievement (Anastasi, 1968).

Although many educators have argued that the above factors are important considerations in interpreting test results (see Sattler & Theye, 1967, for a review), there is little empirical evidence regarding standardized group achievement tests which can be used to estimate the impact of these variables on test results and on subsequent placement and curriculum decisions. Results from studies on I.Q. testing indicate that students' test scores can be raised or lowered by various testing practices (Kirkland, 1971). Familiarity with the test (practice) (Millman, Bishop, & Ebel, 1965), low anxiety levels (Allen, 1972), use of reinforcement (Ayllon & Kelly, 1972; Edlund, 1972; Tiber & Kennedy, 1964), rapport with the tester (Sacks, 1952), and an understanding of the test format (Fueyo, 1977) tend to increase I. Q. scores. If achievement test scores can be raised by motivating and

[1]This research was supported by Handicapped Student Research Grant # 13-443A from the Bureau of Education for the Handicapped.

training students, and training administrators, then the use of standardized tests in Title I programs under "typical conditions" (i.e., untrained students and administrators, and unmotivated students) may be inappropriate and invalid.

Objectives

The purpose of the study was to examine the effects of using group standardized tests to measure the academic achievement of students in Title I classrooms under training and reinforcement conditions. Specific questions addressed are listed below:

1. What is the effect of providing contingent reinforcement to students on reading test scores and on student testing behavior?
2. What is the effect of training teachers in test administration procedures on the testing behavior of teachers, on students' reading test scores, and on student testing behavior?
3. What is the effect of training students in test taking skills on students' reading test scores and on student testing behavior?

Method

Subjects

Twelve Title I teachers, 24 Title I aides, 12 regular classroom teachers, and 597 second grade students in 12 Title I or similar schools (24 classrooms) participated in the study. The effects of reinforcing higher scores and training students and teachers in proper testing procedures were examined during the district's spring administration of the Stanford Achievement Test (SAT, Primary II, Form A). A three-way factorial design was used in which students were randomly assigned by classroom to experimental and control groups of all combinations of independent variables: (a) reinforcing students, (b) training teachers, and (c) training students.

Independent Variables

Reinforcement (RE). Reinforcement was delivered to students based on their performance on the Reading B and Word Study subtests of the SAT. Both subtests were administered the same day and the order of administration was randomly assigned.

One half of the classrooms were randomly selected for reinforcement and students (n = 297) were paid five cents for each point they scored on the post test above their expected score (predicted from the pretest). The subtests were hand scored immediately following test administration and students were reinforced on the same day. Students in the control group received no payment for reading scores, but were reinforced for math scores after reading data was collected.

Teacher training (TT). Twelve Title I teachers and twelve Title I aides were instructed in appropriate test administration techniques. Training consisted of four sections: (a) preparation of students, (b) techniques for group testing, (c) observation of video taped classroom testing scenes, and (d) rehearsal of the SAT administration. Trained teachers and aides were randomly assigned to half of the 24 classrooms and administered the SAT to the students (n = 307). The regular classroom in the remaining 12 classrooms and 12 Title I aides were untrained and administered the test in control group classrooms.

Student training (ST). The students in a random half of the classrooms were trained by the principal investigator for one hour in test taking skills up to two weeks before the actual SAT testing. Skills taught included filling in machine scorable forms, attending to and following directions, analyzing questions and answers, and responding to unique test formats. Students in the control group received no training prior to the test.

Dependent Variables

The effects of providing reinforcement, training teachers, and training students were observed by measuring three variables:

1. Students' scores on the SAT (Reading B and Word Study),
2. Percent of student on-task behavior during the test, and
3. Percent of teacher on-task behavior during the test.

Results

Analysis of variance was performed on each of the dependent measures using classroom as the unit of analysis. The results of the analyses are presented in Table 1 and the marginal means in Table 2. Student scores differed significantly under RE ($p < .001$) and under TT ($p < .09$) conditions. Reinforced student scores ($\bar{X}_{R}$ = 87.5)

were over one standard deviation higher than unreinforced scores ($\overline{X}_{-RE}$ = 77.5) and the scores in trained teacher rooms ($\overline{X}_{TT}$ = 84.8) were 1/2 standard deviation higher than in untrained teacher rooms ($\overline{X}_{-TT}$ = 80.3).

Table 1

Results of Analyses of Variance
Performed for Each Dependent Variable

Source	df	Reading Scores		Teacher On-task		Student On-task	
		MS	F	MS	F	MS	F
Reinforcement	1	598.90	16.8[a]	1594.44	1.00	618.89	1.81
Trained Teachers	1	124.17	3.5[c]	49003.04	30.59[a]	167.48	.49
Trained Students	1	62.05	1.7	442.72	.28	69.55	.20
RE · TT	1	37.93	1.1	374.17	.23	432.55	1.27
RE · TS	1	28.76	.8	13.87	.01	152.65	.45
TT · TS	1	154.79	4.3[b]	.03	.00	22.33	.07
RE · TT · TS	1	10.92	.3	28.11	.02	337.08	.99
Error	16	35.71		1601.98		341.46	
Total	23	69.08		3527.54		310.99	

[a] $p < .001$ [b] $p < .06$ [c] $p < .09$

Table 2

Mean Score and Standard Deviation for Classrooms
Assigned to Each Treatment Group

Treatment	Reading Score		% Teacher On-task		% Student On-task	
	$\overline{X}$	SD	$\overline{X}$	SD	$\overline{X}$	SD
Reinforced	87.5	7.4	54.25	28.95	74.20	6.69
Unreinforced	77.5	5.9	45.21	31.37	70.43	7.67
Trained teachers	84.8	9.1	73.74	13.46	74.64	5.00
Untrained teachers	80.3	7.1	26.55	21.11	71.88	11.58
Trained students	80.9	9.1	53.08	28.89	72.57	10.35
Untrained students	84.1	7.5	47.70	31.41	73.83	7.75

A significant interaction ($p < .06$) between TT and TS treatments was observed. Trained students received higher scores in untrained teacher rooms ($\bar{X}_{-TS} = 81.2$) than in trained teacher rooms ($\bar{X}_{TS} = 80.7$), and untrained students had higher scores in trained teacher rooms ($\bar{X}_{TS} = 89.0$) than in untrained teacher rooms ($\bar{X}_{-TS} = 79.3$).

A significant difference ($p < .001$) in teacher on-task behavior was found between trained teacher rooms ($\bar{X}_{TT} = 73.74$) and untrained teacher rooms ($\bar{X}_{-TT} = 26.55$), indicating that behaviors associated with standardized procedures occurred more frequently with trained teachers than with untrained teachers. No significant differences in student on-task behavior occurred across treatments.

Conclusions

The major finding of the study was that classes producing the highest reading test scores were those associated with trained test administrators and reinforced students. These results demonstrate that group achievement test scores may be raised substantially by ensuring that test administration conditions comply with standardized procedures and that students try to do their best while being tested. Such findings imply that achievement test scores obtained when standardized administration procedures are not followed closely or when students are not motivated to do their best may be invalid estimates of actual academic knowledge and invalid bases for critical placement decisions. In the future, test scores may need to be interpreted in conjunction with the administration condition.

The results of the TT X TS interaction imply that trained teachers had more influence over the untrained students than over the trained students. Since second grade students frequently copy the behavior of their teachers and since, in this study, students were trained by someone other than the test administrators, it is possible that the children became confused as to what was the appropriate behavior.

The nonsignificant differences between trained students and untrained students may lead one to conclude that increased test scores would not be a likely result of training students in test wiseness skills. However, due to the nature of the training conducted in the present study, such a conclusion is not warranted. The

one hour of training provided two weeks before the test by a stranger who was not the test administrator simply was not adequate to change student testing behavior sifficiently to enable them to obtain higher test scores. Furthermore, the lack of significant differences in on-task behavior between the trained and untrained students supports the notion that training as implemented in this study was not effective.

Since decisions in the areas of student selection, placement, curriculum, funding, and evaluation are often based on group standardized achievement test data, it is critical that variables other than student achievement be controlled across testing sites. Given that test scores are such a pervasive and important part of our educational system, the results of this study indicate that more effort should be devoted to ensuring that tests are properly administered and students are properly motivated.

The results of this study should not be used to support the notion that students be paid money to get high scores, but should suggest that Title I students may not be scoring at their true level of potential because of low motivation. Further research is clearly needed to determine the appropriate techniques for reinforcing students to try their best during testing and still have standardized procedures maintained throughout.

References

Allen, G. J. The behavioral treatment of test anxiety: Recent research and future trends. Behavior Therapy, 1972, 3, 253-363.

Anastasi, A. Psychological testing. New York: MacMillan, 1968.

Ayllon, T., & Kelly, K. Effects of reinforcement on standardized test Performance. Journal of Applied Behavior Analysis, 1972, 5, 477-484.

Edlund, C. The effect on the behavior of children, as reflected in the I.Q. scores when reinforced after each correct response. Journal of Applied Behavior Analysis, 1972, 5, 317-319.

Fueyo, V. Training test-taking skills: A critical analysis. Psychology in the Schools, 1977, 14, 180-184.

Kirkland, M. C. The effects of tests on students and schools. Review of Educational Research, 1971, 41, 303-349.

Millman, J., Bishop, C. H., & Ebel, R. An analysis of test-wiseness. Educational and Psychological Measurement, 1965, 25, 707-726.

Piersel, W. C., Brody, G. H., & Kratochwill, T. R. A further examination of motivational influences on disadvantaged minority group children's intelligence test performance. Child Development, 1977, 48, 1142-1145.

Sacks, E. L. Intelligence scores as a function of experimentally established social relationships between child and examiner. Journal of Abnormal and Social Psychology, 1952, 47, 354-358.

Sattler, J. M., & Theye, R. Procedural, situational, and interpersonal variables in individual intelligence testing. Psychological Bulletin, 1967, 68, 347-360.

Tiber, N., & Kennedy, W. A. The effects of incentives on the intelligence test performance of different social groups. Journal of Consulting Psychology, 1964, 28, 187.